A Note on the Strengths and Weaknessess of Using the CPS To Estimate Children's Health Insurance Coverage

Katherine Swartz

The Urban Institute

Washington, DC

February 1988

Contractor Document
Health Program, Office of Technology Assessment
U.S. Congress, Washington, DC 20510

This paper was prepared by an outside contractor for the OTA assessment Healthy Children: Investing in the Future. The paper does not necessarily reflect the analytical findings of OTA, the assessment's advisory panel, or the Technology Assessment Board.

J952-13

A NOTE ON THE STRENGTHS AND WEAKNESSES OF USING THE CPS TO ESTIMATE CHILDREN'S HEALTH INSURANCE COVERAGE

Introduction[1]

This paper discusses the strengths and weaknesses of the Current Population Survey (CPS) relative to the National Health Interview Survey (NHIS), the 1977 National Medical Care Expenditure Survey (NMCES), and the 1980 National Medical Care Utilization and Expenditure Survey (NMCUES), for the purpose of estimating the numbers of children with different types of health insurance. CPS has asked about health insurance coverage in every March since 1979, and the NHIS asks about health insurance in even numbered years. The U.S. Bureau of the Census conducts the surveys for the CPS and NHIS; the CPS data are edited and made available for public use by the Census, while the NHIS data are edited and made available by the National Center for Health Statistics. The NMCES and NMCUES surveys were both conducted jointly by Research Triangle Institute and the National Opinion Research Corporation. The National Center for Health Services Research--now called the National Center for Health Services Research and Health Care Technology Assessment--analyzed the NMCES data, while the National Center for Health Statistics edited and made available the NMCUES data.

---

1 Much of the material in this paper draws upon earlier research done by the author (see especially 23.) She is particularly grateful to John Coder and Chuck Nelson of the Bureau of the Census, and to Patricia Ruggles and Sheila Zedlewski of The Urban Institute for further discussions about the CPS. Any errors of interpretation are, of course, the author's sole responsibility.

Reasons for Differences and Similarities Across Surveys

In this section we examine four factors which might contribute to differences and similarities in estimates from different surveys, and which might make the CPS relatively strong or weak for purposes of this project. The four factors are: the sample framework design, the weighting of the survey responses, the handling of nonresponses to questions, and the wording of questions.

Sample Framework Design

The four surveys which we are reviewing in this paper all have samples composed of people who live in housing units which were selected for the respective surveys. Hence, the focus of the four sample framework designs is the housing unit and not a particular person or family.[2] What emerges from the literature on the sample designs for the four surveys is that they are very similar. They are all probability samples. The major difference between them is the final number of households assigned for interview (see table 1). The CPS and NHIS samples are clearly far larger than those of the NMCES and NMCUES, and they are therefore able to include housing units from more areas of the United States. The CPS contains almost 58,000 children--a number which is about one and a half times the size of the whole NMCES sample. Larger samples always produce more precise estimates--a clear advantage of the CPS and NHIS. The widely held opinion is that the sample sizes probably do not contribute to the differences in the estimates of the aggregate numbers of

---

2 This type of design differs from others such as a telephone survey. In a telephone survey the focus of the sample design is the telephone number, which typically is accessed via random digit dialing.

Table 1

Numbers of Primary Sampling Units and Housing Units in the Four Surveys

	NMCES	1978 HIS	March 1980 CPS	NMCUES*
Number of housing units assigned for interview	16,684	51,000	80,500	8,000
Non-interview rate (type A non-response at 1st round interview)	8.4%	3.8%	4.3%	8.9%
Number of people in the sample (all ages)	38,815	110,000	181,488	17,123

*The figures for NMCUES relate only to the national sample and do not include the sample of Medicaid households from New York, California, Michigan, and Texas.

people with and without health insurance.[3] However, estimates of the numbers of people in subgroups of the population, such as children in a particular age cohort with and without health insurance, are undoubtedly more reliable from the CPS because it has a larger sample size.

Weighting of Survey Responses

The weighting procedures used for all four surveys are for all practical purposes identical even though they differ in some details.[4] Thus, it cannot be argued that the weighting procedure used by one or another of the four surveys make any one of them relatively stronger or weaker.

Nonresponses

A third factor which might cause surveys to differ and which might make one survey relatively strong or weak is the way in which nonresponses to various questions are handled.

The extent to which item nonresponses are a problem is not publicly available for any of the surveys except the CPS. The CPS nonresponse rates for the questions related to health insurance are between 10 and 15.5 percent. Depending on which of the health insurance questions has a nonresponse, the CPS uses one of three hot-deck procedures to impute a missing answer. The 1978 NHIS nonresponses were given answers on the basis of earlier answers and notes which the interviewer wrote on the questionnaire answer form. So-called decision logic tables were created to give data coders

---

3 Not one of the people I have talked with in connection to this problem felt it would account for more than a fraction of a percent of the differences in the estimates of the proportion of uninsured people.

4 See Swartz (23) for the details.

instructions on how missing answers should be filled in. Since the procedures for filling in missing answers on the NHIS are so vague, we cannot compare them with the imputation procedures used in the other three surveys.

NMCES and NMCUES have a compounded problem in dealing with nonresponse items because of their longitudinal nature. Not only were there item nonresponses for respondents but there was also attrition from the sample. In the case of NMCES, 10.7 percent of the participants from the first interview failed to respond for at least one other interview during the time they were eligible.[5] NMCES retained in the sample anyone who responded to the initial interview--even if they failed to respond to the other four interviews. NMCUES retained in the sample anyone who responded to interviews during one-third or more of the time for which he or she was eligible. For both NMCES and NMCUES, the person was assumed to have the same health insur-ance at the missing interviews as he or she had when responding. This strategy of looking at the person's answers in other interview rounds was also used to allocate nonresponses to the health insurance questions when the respondent answered all of the interviews but did not respond to all of the insurance questions.

There are potential problems with the imputation methods used for the CPS, NMCES, and NMCUES. The most serious problem applies to all three surveys. The imputation methods assume that whether or not an observation is missing is independent of its true value. But, of course, when this assumption is not valid, the imputed estimates will be biased. In the case of health insurance, this assumption is tenuous, but the degree to which each of the survey's allocated answers are biased because of it is unknown.

---

5 People who died or were born during the life of the survey obviously could not answer all five interviews. They are therefore considered eligible only for interviews during the time period in which they were alive.

In reviewing the imputation procedures for the CPS, it is important to remember that we are not able to apply the same scrutiny to the imputation procedures for the NHIS, NMCES, and NMCUES. Thus, what follows may appear to cast doubts on the validity of the CPS; we suspect, however, that the other surveys have similar problems with imputing for nonresponses and under-reporting. Income sources and amounts, for example, are frequently under-reported on household surveys. In the case of the March 1984 CPS, only 66 percent of the total dollars that the Office of Family Assistance believes was paid to families receiving Aid to Families With Dependent Children (AFDC) was reported by the survey respondents.[6] The Census has adopted a conservative approach for imputing who gets income from different sources and how much of that type of income is received--the approach is conservative in that not all of the difference between what is reported to the CPS and what is independently estimated to be the actual total is allocated to the survey respondents. In the case of AFDC income, the Census only allocated another 9 percent of the Office of Family Assistance total to families. Hence, the Census allocation procedure not only does not allocate all of the difference in dollars but it also must be missing some families who in fact received AFDC cash assistance.

We dwell on the AFDC example because the imputation procedure for who receives Medicaid uses the AFDC responses (both reported and allocated). If a family has a positive, reported response for AFDC income, the Medicaid answer is checked to make sure that it, too, is positive; if by chance it is negative, the editing routine changes the answer to a positive response. If the family is allocated AFDC income, the same procedure automatically

---

6 See (24), Table A-2, p. 170.

allocates a positive response for Medicaid recipiency. Thus, the CPS problem with nonreporting of AFDC feeds into the underreporting for Medicaid. The nonreporting of Supplemental Security Income income similarly contributes to underreporting of Medicaid.

Our best estimate of the undercounting of all Medicaid recipients on the March 1984 CPS is about 550,000. The Social Security Bulletin reported an unduplicated number of Medicaid recipients in fiscal year 1984 of 21,365,000 (25). The Health Care Financing Administration estimates that 7 percent of Medicaid recipients were institutionalized in 1984, which would be 1,496,000. The noninstitutionalized total would then by 19,869,000. My March 1984 CPS estimate of the number of noninstitutionalized Medicaid recipients is 19,307,000.

Prior to 1982, the CPS problem with underreporting of AFDC and Medicaid was larger because subfamilies headed by an unmarried woman were not identified.[7] That is, an unmarried woman and her child or children were not identified as a sub-family if they lived with the woman's parents. Thus, if AFDC recipiency was not reported, the Census allocation and imputation procedures would not count the sub-family as eligible for AFDC. Since 1982, sub-families headed by unmarried women have been identified.

Question Differences

The fourth factor which may contribute to differences in the estimates from the four surveys concerns the differences in the focus of each survey, question wording, and the ordering of questions. The NHIS, NMCES, and NMCUES

---

7 A preliminary estimate of the number of women and children in such families who are not previously identified as receiving AFDC and Medicaid in 1979 (using the March 1980 CPS) is about 500,000.

are all surveys which focus on health issues. In contrast, the March CPS focuses on the previous year's employment and income; the five major questions related to health insurance are at the very end of the questionnaire. Clearly, if questions related to medical care expenditures and utilization trigger people's memories about their health insurance better than do questions about employment and income, the NHIS, NMCES, and NMCUES have better responses. However, since NMCES data indicate that 84 percent of the adults who were in the labor force had private health insurance, asking about employment and income may do as well in terms of triggering people's memories about their health insurance coverage.

Another issue related to the focus of the surveys concerns the longitudinal nature of the NMCES and NMCUES samples, and the semi-longitudinal nature of the CPS sample.[8] The NMCES and NMCUES sample respondents were asked the same questions about health insurance (as well as other questions) every 3 months over a 14-month period. In the CPS, the health insurance questions are only asked in March every year. But about half of all the sample respondents in March were also sample respondents in the previous March, and may remember that questions related to health insurance were asked the year before.

A type of bias long known as rotation-group bias can occur in people's answers when they are in longitudinal samples.[9] The answers given by people when they are first interviewed often differ markedly from their answers in subsequent interviews. The problem caused by rotation-group bias is that we do not know whether the value of the initial response or the value of the

---

8 A longitudinal sample is one where the sample respondents are resurveyed one or more times. The HIS sample is not a longitudinal sample--every year HIS interviews a different sample of people.

9 See Bailar (1).

subsequent responses is closer to the "true" value. In NMCES, the first and last interviews have the highest proportion of uninsured (13.8 percent) while in the March 1984 CPS, it is the first rotation group that has the lowest proportion of people without health insurance (15.4 percent).[10] Thus, the effects of rotation-group bias on the longitudinal surveys here are difficult to assess. It is likely that the procedures for dealing with nonresponses on the health insurance questions have confounded the usual direction of rotation-group bias.

The questions pertaining to health insurance on the four surveys are worded differently and appear in different orders. The most obvious difference in the wording of the health insurance questions is that the CPS asks about the previous year while the other three surveys ask about the time of the interview. In assessing whether the CPS respondents are answering the health insur-ance questions with respect to the previous year, it is important to recall that the health insurance status of most people does not vary from year to year. Using heads of households 19 to 64 years old on the NMCES public use file, Swartz (22) found that just 8.4 percent were insured only some of the five interviews. Another way of looking at this is that 95 percent of NMCES household heads 19 to 64 years old who had health insurance in the first interview had insurance in all five interviews. Similarly, 70 percent of those heads who had no coverage in the first interview had no coverage in all five interviews. Thus, the memory recall problem for the CPS respondents really exists only for the small part of the sample who had coverage during some part but not all of the previous year.

The four surveys' questions on Medicaid differ, too. NMCES asks about

---

10 See Swartz (22) for more details.

Medicaid and other public assistance in the same question while NMCUES separates the two, and both the CPS and NHIS never ask about coverage through other public assistance programs that pay part of a person's medical bills. Only the CPS does not ask to see a person's Medicaid card. Overall, it is difficult to determine a quantitative estimate of the impact of the question wording differences.

It is not clear that the differences in the health insurance question ordering among the surveys should have a substantive impact. While the CPS health insurance questions are the last questions on the questionnaire, and may therefore elicit responses from tired respondents, the other surveys place their health insurance questions almost at the end of the questionnaires. Perhaps more importantly, the NHIS, NMCES, and NMCUES are longer questionnaires than the CPS so the issue of respondents being tired of answering questions is the same for all of the surveys.

Finally, children who are younger than 15 years old are not directly the subject of questions about health insurance coverage in the CPS, whereas they are for the NMCES and NMCUES surveys. For almost all children, this is not a problem because the CPS editing allocates a yes or no response for employer-group and self-purchase/other health insurance if the parent says he or she has a policy which also covers the children. In the case of Medicare, Medicaid, and CHAMPUS/VA (Civilian Health and Medical Program of the Uniformed Services/Veterans Administration) coverage, the CPS interviewer asks if anyone (regardless of age) in the family is covered by or has had medical bills paid by each of the three Government-provided programs. Thus, the only children who are potentially incorrectly counted as not having health insurance are those who are covered by a policy bought by a parent who does not live in the household.

We can estimate roughly how many children are in this position by using the April 1984 CPS.[11] The April 1984 CPS had a supplemental questionnaire for women with children under 21 years of age from absent fathers. On the basis of this survey, the Bureau of the Census estimates that there were 8,690,000 women (plus or minus 190,000) who had custody of children under the age of 21 from absent fathers. A quarter of these women had married again, and would be indistinguishable from women who had been married only once on the March CPS--i.e., their marital status in March would be "married." Among all the women with children from an absent father, 3,995,000 (46 percent) were supposed to receive child support payments in 1983, but only 3,037,000 (35 percent) actually received the payments. Thus, just over a third of all women with children under 21 years of age from an absent father received child support payments in 1983. When we look at whether or not health insurance was included in the child support award or agreement, we find that a little over half of the women who actually received child support payments also had health insurance included (1,641,000). On average, these women had about 1.8 children, or about 2,954,000 children under 21 years of age altogether. In 1984, 59.1 percent of all children under 21 were 12 years old or younger, so, for this age group, about 1,746,000 children had health insurance from absent parents. An unknown percentage of these children would have been reported as uninsured, when in fact they have insurance through their child custody agreements. If the child had no other source of insurance but that of the absent parent, then the CPS would have shown him or her to be uninsured.

Based on the information above, we can calculate a lower bound on the

---

11 The information that follows comes from "Child Support and Alimony: 1983 (Supplemental Report)," Current Population Reports, Special Studies, Series P-13, No. 148, by the U.S. Department of Commerce, Bureau of the Census, 1986 (8).

proportion of children ages 0-12 who were without insurance. From the March 1984 CPS, we estimated there were 7,873,000 children between 0 and 12 years of age who were uninsured (18 percent of all children 0-12 years of age). If we assume that none of the children with insurance through child custody agreements had other sources of insurance, then the actual number of uninsured children would be 6,127,000 (7,873,000 minus 1,746,000), which is 14 percent of the total 0-12-year-old population in 1984. Thus, the lower bound estimate of the number of uninsured children 0 to 12 years old in 1984 is 78 percent of the reported number.

Applying this rate (78 percent) to the March 1986 CPS estimate of the number of uninsured children also yields a lower bound estimate of 14 percent. Thus, the true proportion of uninsured children between 0 and 12 years of age in 1986 lies somewhere in the range of 14 to 18 percent.

It is also possible to calculate a lower bound on the proportion of children with divorced, separated, or never married custodial parent and an absent parent who were without insurance. The 8,690,000 women mentioned above had an average of 1.8 children per woman, or about 15,642,000 children under 21 years of age altogether. Since the March CPS does not allow us to distinguish between women who have been married only once and women who have remarried, we need to remove remarried women with children from an absent father from this total. There were 2,129,000 remarried women and 3,832,000 children (at 1.8 children per woman). Thus, there were 11,810,000 children younger than 21 who lived with a divorced, separated, remarried but widowed from a later marriage, or never married mother and an absent father. Of the 1,641,000 women who received child support payments in 1983 and had health insurance included in the award, 29 percent or 476,000 were remarried. Thus,

widowed, or never married) had approximately 2,097,000 children. These children with health insurance included in child support agreements represent 17.8 percent of all children of divorced, separated, remarried but widowed, or never married women.

From the March 1984 CPS, we estimated there were 4,547,000 uninsured children under 18 years of age who had a divorced, separated, or never married custodial parent and an absent parent. If all of the children who had insurance through the absent parent's policy were reported as uninsured in the CPS, then the actual number of uninsured children of custodial parents in these marital categories would be only 62 percent of the reported number.[12] Applying this rate to the 1986 estimate of the number of uninsured children we find an overall percent uninsured of 22 percent. This rate of uninsuredness is clearly a lower bound: the true rate lies somewhere in the range of 22 percent to 35 percent.

CONCLUSIONS

The strengths of the CPS for estimating the health insurance coverage of children are that it has a very large sample (almost 58,000 children) and that it is conducted every year so it is the most recent survey of health insurance coverage in the United States. The weaknesses of the CPS for this project are that it underestimates the number of children with Medicaid coverage and the number of children of divorced, separated, or never-married parents with health insurance provided by an absent parent. We do not know the number of children who actually have Medicaid or private health insurance but appear to be uninsured. My own estimate is that the number is, at most, one million.

---

12 Calculated as: (4,547,000 - (2,097,000 x .83)) / 4547000, where 83 is the percent of children between 0-20 who were 0 to 17 years old.

In evaluating these strengths and weaknesses, estimates of the numbers of children in various subgroups (e.g., by age cohort, race, health insurance type, and marital status of the mother) could not be obtained with the NMCES or NMCUES because their sample sizes would not support such groupings. The CPS has a clear and large advantage with its sample size for estimating such sub-groups. Given this advantage, the weakness of the CPS in terms of underestimating children with Medicaid coverage or with private health insurance paid for by an absent parent seems outweighed. The other three surveys undoubtedly also have such problems with underestimating people with particular types of health insurance; we know of the problems with the CPS because of its greater use. Finally, the fact that the CPS allows us to estimate the numbers of children in these subgroups in 1984, and in a few months for 1986, is a clear boon for policy debates.

REFERENCES

1. Bailar, B.R., "The Effects of Rotation Group Bias on Estimates from Panel Surveys," Journal of the American Statistical Association, 70(349):23-30, March 1975.

2. Bailar, J.C., III, and Bailar, B.A., "Comparison of Two Procedures for Imputing Missing Survey Values." American Statistical Association 1978 Proceedings of the Section on Survey Research Methods, pp. 462-467, Aug. 14-17, 1978.

3. Berk, M.L., Wilensky, G.R., and Cohen, S.B., "Methodological Issues in Health Surveys: An Evaluation of Procedures Used in the National Medical Care Expenditure Survey." Evaluation Review, Vol. 8, No. 3, June 1984.

4. Bonham, G.S., "Procedures and Questionnaires of the National Medical Care Utilization and Expenditure Survey." Series A, Methodological Report No. 1, U.S. Department of Health and Human Services Publication No. 83-20001, National Center for Health Statistics, March 1983.

5. Bonham, G.S., and L.S. Corder, "NMCES Instruments and Procedures 1: Household Interview Instruments." U.S. Department of Health and Human Services Publication No. (PHS) 81-3280, National Center for Health Services Research, April 1981.

6. U.S. Department of Commerce, Bureau of the Census, "The Current Population Survey: Design and Methodology." Technical Paper 40, January 1978.

7. U.S. Department of Commerce, Bureau of the Census, "Workbook for Interviewers." Current Population Survey Group Training, March 1984.

8. U.S. Department of Commerce, Bureau of the Census, "Child Support and Alimony: 1983 (Supplemental Report)," Current Population Reports, Special Studies, Series P-13, No. 148, 1986.

9. Casady, R.J. "The Survey Design and Estimation Methodology for the National Medical Care Utilization and Expenditure Survey," Proceedings of the 19th National Meeting of the Public Health Conference on Records and Statistics, August 1983, DHHS Pub. No. (PHS) 84-1214, National Center for Health Statistics, December 1983.

10. Cohen, S.B. "An Analysis of Alternative Imputation Strategies for Individuals With Partial Data in the National Medical Care Expenditure Survey." Review of Public Data Use, Vol. 10, 1982, pp. 153-165.

11. Cohen, S.B., and Farley, P.J., "NMCES Instruments and Procedures 3: Estimation and Sampling Procedures in the NMCES Insurance Surveys." U.S. Department of Health and Human Services Publication No. (PHS) 84-3369, National Center for Health Services Research, May 1984.

12. Cohen, S.B., and Kalsbeek, W.D., "NMCES Instruments and Procedures 2: Estimation and Sampling Variances in the Household Survey." U.S. Department of Health and Human Services Publication No. (PHS) 81-3281,

National Center for Health Services Research, June 1981.

13. Cox, B.G., "Compensating for Missing Data in the NMCUES," Proceedings of the 19th National Meeting of the Public Health Conference on Records and Statistics, August 1983, DHHS Pub. No. (PHS) 84-1214, National Center for Health Statistics, December 1983.

14. Dicker, M., "Health Care Coverage and Insurance Premiums of Families: United States, 1980." National Medical Care Utilization and Expenditure Survey Preliminary Data Report No. 3, U.S. Department of Health and Human Services Publication No. 83-20000, National Center for Health Statistics, May 1983.

15. Kasper, J.A., Walden, D.C., and Wilensky, G.R., "Who Are the Uninsured?" National Health Care Expenditures Study Data Preview 1, National Center for Health Services Research, November 1980.

16. Kerachsky, S.H., and Mallar, C.D.. "The Effects of Monetary Payments on Survey Responses: Experimental Evidence from a Longitudinal Study of Economically Disadvantaged Youths." American Statistical Association 1981 Proceedings of the Section on Survey Research Methods, Aug. 10-13, 1981.

17. Moore, T.F., Bettin, P., Kostanich, D., et al., "Overview of Current Population Survey Sample Design." American Statistical Association 1979 Proceedings of the Section on Survey Research Methods, Aug. 13-16, 1979.

18. U.S. Department of Health, Education, and Welfare, Public Health Service, National Center for Health Statistics: "Current Estimates From the Health Interview Survey: United States 1978", series 10-No. 130, U.S. DHEW Pub. No. (PHS) 80-1551, November 1979.

19. Piper, L.L., "NMCUES Household Survey Sample Design Statement", Research Triangle Institute Project No. 25IU-1895-5, Working Paper No. 1 (Draft), December 1979.

20. Schooley, M., Cahoon, L., Ponikowski, C., et al., "Overview of the Current Population Survey Operations and Estimation." American Statistical Association 1979 Proceedings of the Section on Survey Research Methods, Aug. 13-16, 1979.

21. Schuman, H., and Presser, S., Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Context (New York: Academic Press, Inc., 1981).

22. Swartz, K., "Who Has Been Without Health Insurance? Changes Between 1963 and 1979." Urban Institute Working Paper 3308-01, Washington, DC, April 1984.

23. Swartz, K., "How Different Are Four Surveys' Estimates of the Number of Americans Without Health Insurance?," Urban Institute Working Paper 3442-01, Washington, DC, December 1984.

24. U.S. Department of Commerce, Bureau of the Census. Money Income of Households, Families, and Persons in the United States: 1984. Current Population Reports, Consumer Income Series P-60, No. 151, April 1986.

25. U.S. Department of Health and Human Services, Social Security Administration, Social Security Bulletin: Annual Statistical Supplement, 1984-85. SSA. Pub. No. 13-11700, Washington, DC.

26. Walden, D.C., Horgan, C.M., and Cafferata, G.L., "Consumer Knowledge of Health Insurance Coverage," prepared for the Fourth Conference on Health Survey Research Methods, Washington, DC, May 3-5, 1982.

27. Walden, D.C., Wilensky, G.R., and Kasper, J.A., "Persons With Changing Insurance Status in 1977." National Health Care Expenditures Study Data Preview 21, draft mimeo, National Center for Health Services Research, August 1984.

28. Wilensky, G.R., Walden, D.C., and Kasper, J.A. "The Uninsured and Their Use of Health Services," presented at the Annual Meetings of the American Statistical Association, August 1981.

29. Williams, R.L., and Folsom, R.E., "Weighted Hot-Deck Imputation of Medical Expenditures Based on a Record Check Subsample," American Statistical Association 1981 Proceedings of the Section on Survey Research Methods, Aug. 10-13, 1981.

30. Wright, R.A., "Survey Design and Administration of the National Medical Care Utilization and Expenditure Survey," Proceedings of the 19th National Meeting of the Public Health Conference on Records and Statistics, August 1983, DHHS Pub. No. (PHS) 84-1214, National Center for Health Statistics, December 1983.